

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Robertson et al.
- (ii) TITLE OF INVENTION: Catalases
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: WORD PERFECT 5.1
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: 08/674,887
- (B) FILING DATE: July 3, 1996
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Charles J. Herron
- (B) REGISTRATION NUMBER: 28,019
- (C) REFERENCE/DOCKET NUMBER: 331400-55
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 201-994-1700
- (B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 52 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAATAAC GCATCCGCTG AC

52

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 31 NUCLEIC ACID
- (B) TYPE: NUCLEOTIDES
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAAGCTT TTACGACGCG ACGTCGAAAC G

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGGAAAAT CACAAACACT CA

52

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGAAGGTACC TTATTTTCAGA TCAAACCGGT C

31

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 2262 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG AAT AAC GCA TCC GCT GAC GAT CTA CAC AGT AGC TTG CAG CAA AGA
Met Asn Asn Ala Ser Ala Asp Asp Leu His Ser Ser Leu Gln Gln Arg
5 10 15

48

TGC AGA GCA TTT GTT CCC TTG GTA TCG CCA AGG CAT AGA GCA ATA AGG
Cys Arg Ala Phe Val Pro Leu Val Ser Pro Arg His Arg Ala Ile Arg
20 25 30

96

GAG AGA GCT ATG AGC GGT AAA TGT CCT GTC ATG CAC GGT GGT AAC ACC
Glu Arg Ala Met Ser Gly Lys Cys Pro Val Met His Gly Gly Asn Thr
35 40 45

144

TCG ACC GGT ACT TCC AAC AAA GAT TGG TGG CCG GAA GGG TTG AAC CTG Ser Thr Gly Thr Ser Asn Lys Asp Trp Trp Pro Glu Gly Leu Asn Leu 50 55 60	192
GAT ATT TTG CAT CAG CAA GAT CGC AAA TCA GAC CCG ATG GAT CCG GAT Asp Ile Leu His Gln Gln Asp Arg Lys Ser Asp Pro Met Asp Pro Asp 65 70 75 80	240
TTC AAC TAC CGT GAA GAA GTA CGC AAG CTC GAT TTC GAC GCG CTG AAG Phe Asn Tyr Arg Glu Glu Val Arg Lys Leu Asp Phe Asp Ala Leu Lys 85 90 95	288
AAA GAT GTC CAC GCG TTG ATG ACC GAT AGC CAA GAG TGG TGG CCC GCT Lys Asp Val His Ala Leu Met Thr Asp Ser Gln Glu Trp Trp Pro Ala 100 105 110	336
GAC TGG GGG CAC TAC GGC GGT TTG ATG ATC CGT ATG GCT TGG CAC TCC Asp Trp Gly His Tyr Gly Gly Leu Met Ile Arg Met Ala Trp His Ser 115 120 125	384
GCT GGC ACC TAC CGT ATT GCT GAT GGC CGT GGG GGC GGT GGT ACC GGA Ala Gly Thr Tyr Arg Ile Ala Asp Gly Arg Gly Gly Gly Thr Gly 130 135 140	432
AGC CAG CGC TTT GCA CCG CTC AAC TCC TGG CCG GAC AAC GTC AGC CTG Ser Gln Arg Phe Ala Pro Leu Asn Ser Trp Pro Asp Asn Val Ser Leu 145 150 155 160	480
GAT AAA GCG CGC CGT CTG CTG TGG CCG ATC AAG AAG AAG TAC GGC AAC Asp Lys Ala Arg Arg Leu Leu Trp Pro Ile Lys Lys Lys Tyr Gly Asn 165 170 175	528
AAA ATC AGC TGG GCA GAC CTG ATG ATT CTG GCT GGC ACC GTG GCT TAT Lys Ile Ser Trp Ala Asp Leu Met Ile Leu Ala Gly Thr Val Ala Tyr 180 185 190	576
GAG TCC ATG GGC TTA CCT GCT TAC GGC TTC TCT TTC GGC CGC GTC GAT Glu Ser Met Gly Leu Pro Ala Tyr Gly Phe Ser Phe Gly Arg Val Asp 195 200 205	624
ATT TGG GAA CCC GAA AAA GAT ATC TAC TGG GGT GAC GAA AAA GAG TGG Ile Trp Glu Pro Glu Lys Asp Ile Tyr Trp Gly Asp Glu Lys Glu Trp 210 215 220	672
CTG GCA CCT TCT GAC GAA CGC TAC GGC GAC GTG AAC AAG CCA GAG ACC Leu Ala Pro Ser Asp Glu Arg Tyr Gly Asp Val Asn Lys Pro Glu Thr 225 230 235 240	720
ATG GAA AAC CCG CTG GCG GCT GTC CAA ATG GGT CTG ATC TAT GTG AAC Met Glu Asn Pro Leu Ala Ala Val Gln Met Gly Leu Ile Tyr Val Asn 245 250 255	768
CCG GAA GGT GTT AAC GGC CAC CCT GAT CCG CTG AGA ACC GCA CAG CAG Pro Glu Gly Val Asn Gly His Pro Asp Pro Leu Arg Thr Ala Gln Gln 260 265 270	816
GTA CTT GAA ACC TTC GCC CGT ATG GCG ATG AAC GAC GAA AAA ACC GCA Val Leu Glu Thr Phe Ala Arg Met Ala Met Asn Asp Glu Lys Thr Ala 275 280 285	864
GCC CTC ACA GCT GGC GGC CAC ACC GTC GGT AAT TGT CAC GGT AAT GGC Ala Leu Thr Ala Gly Gly His Thr Val Gly Asn Cys His Gly Asn Gly 290 295 300	912
AAT GCC TCT GCG TTA GCC CCT GAC CCA AAA GCC TCT GAC GTT GAA AAC Asn Ala Ser Ala Leu Ala Pro Asp Pro Lys Ala Ser Asp Val Glu Asn 305 310 315 320	960

CAG GGC TTA GGT TGG GGC AAC CCC AAC ATG CAG GGC AAG GCA AGC AAC Gln Gly Leu Gly Trp Gly Asn Pro Asn Met Gln Gly Lys Ala Ser Asn 325 330 335	1008
GCC GTG ACC TCG GGT ATC GAA GGT GCT TGG ACC ACC AAC CCC ACG AAA Ala Val Thr Ser Gly Ile Glu Gly Ala Trp Thr Thr Asn Pro Thr Lys 340 345 350	1056
TTC GAT ATG GGC TAT TTC GAC CTG CTG TTC GGC TAC AAT TGG GAA CTG Phe Asp Met Gly Tyr Phe Asp Leu Leu Phe Gly Tyr Asn Trp Glu Leu 355 360 365	1104
AAA AAG AGT CCT GCC GGT GCC CAC CAT TGG GAA CCG ATT GAC ATC AAA Lys Lys Ser Pro Ala Gly Ala His His Trp Glu Pro Ile Asp Ile Lys 370 375 380	1152
AAG GAA AAC AAG CCG GTT GAC GCC AGC GAC CCC TCT ATT CGC CAC AAC Lys Glu Asn Lys Pro Val Asp Ala Ser Asp Pro Ser Ile Arg His Asn 385 390 395 400	1200
CCG ATC ATG ACC GAT GCG GAT ATG GCG ATA AAG GTA AAT CCG ACC TAT Pro Ile Met Thr Asp Ala Asp Met Ala Ile Lys Val Asn Pro Thr Tyr 405 410 415	1248
CGC GCT ATC TGC GAA AAA TTC ATG GCC GAT CCT GAG TAC TTC AAG AAA Arg Ala Ile Cys Glu Lys Phe Met Ala Asp Pro Glu Tyr Phe Lys Lys 420 425 430	1296
ACT TTC GCG AAG GCG TGG TTC AAG CTG ACG CAC CGT GAC CTG GGC CCG Thr Phe Ala Lys Ala Trp Phe Lys Leu Thr His Arg Asp Leu Gly Pro 435 440 445	1344
AAA TCA CGT TAC ATC GGC CCG GAA GTG CCG GCA GAA GAC CTG ATT TGG Lys Ser Arg Tyr Ile Gly Pro Glu Val Pro Ala Glu Asp Leu Ile Trp 450 455 460	1392
GAA GAC CCG ATT CCG GCA GGT AAC ACC GAC TAC TGC GAA GAA GTG GTC Gln Asp Pro Ile Pro Ala Gly Asn Thr Asp Tyr Cys Glu Glu Val Val 465 470 475 480	1440
AAG CAG AAA ATT GCA CAA AGT GGC CTG AGC ATT AGT GAG ATG GTC TCC Lys Gln Lys Ile Ala Gln Ser Gly Leu Ser Ile Ser Glu Met Val Ser 485 490 495	1488
ACC GCT TGG GAC AGT GCC CGT ACT TAT CGC GGT TCC GAT ATG CGC GGC Thr Ala Trp Asp Ser Ala Arg Thr Tyr Arg Gly Ser Asp Met Arg Gly 500 505 510	1536
GGT GCT AAC GGT GCC CGC ATT CGC TTG GCC CCA CAG AAC GAG TGG CAG Gly Ala Asn Gly Ala Arg Ile Arg Leu Ala Pro Gln Asn Glu Trp Gln 515 520 525	1584
GGC AAC GAG CCG GAG CGC CTG GCG AAA GTG CTG AGC GTC TAC GAG CAG Gly Asn Glu Pro Glu Arg Leu Ala Lys Val Leu Ser Val Tyr Glu Gln 530 535 540	1632
ATC TCT GCC GAC ACC GGC GCT AGC ATC GCG GAC GTG ATC GTT CTG GCC Ile Ser Ala Asp Thr Gly Ala Ser Ile Ala Asp Val Ile Val Leu Ala 545 550 555 560	1680
GGT AGC GTA GGC ATC GAG AAA GCC GCG AAA GCA GCA GGT TAC GAT GTG Gly Ser Val Gly Ile Glu Lys Ala Ala Lys Ala Ala Gly Tyr Asp Val 565 570 575	1728
CGC GTT CCC TTC CTG AAA GGC CGT GGC GAT GCG ACC GCC GAG ATG ACC Arg Val Pro Phe Leu Lys Gly Arg Gly Asp Ala Thr Ala Glu Met Thr 580 585 590	1776

GAC GCA GAC TCC TTC GCA CCG CTG GAG CCG CTG GCC GAT GGC TTC CGC Asp Ala Asp Ser Phe Ala Pro Leu Glu Pro Leu Ala Asp Gly Phe Arg 595 600 605	1824
AAC TGG CAG AAG AAA GAG TAT GTG GTG AAG CCG GAA GAG ATG CTG CTG Asn Trp Gln Lys Lys Glu Tyr Val Val Lys Pro Glu Met Leu Leu 610 615 620	1872
GAT CGT GCG CAG CTG ATG GGC TTA ACC GGC CCG GAA ATG ACC GTG CTG Asp Arg Ala Gln Leu Met Gly Leu Thr Gly Pro Glu Met Thr Val Leu 625 630 635 640	1920
CTG GGC GGT ATG CGC GTA CTG GGC ACC AAC TAT GGT GGC ACC AAA CAC Leu Gly Gly Met Arg Val Leu Gly Thr Asn Tyr Gly Gly Thr Lys His 645 650 655	1968
GGC GTA TTC ACC GAT TGT GAA GGC CAG TTG ACC AAC GAC TTT TTT GTG Gly Val Phe Thr Asp Cys Glu Gly Gln Leu Thr Asn Asp Phe Phe Val 655 660 665	2016
AAC CTG ACC GAT ATG GGG AAC AGC TGG AAG CCG GTA GGT AGC AAC GCC Asn Leu Thr Asp Met Gly Asn Ser Trp Lys Pro Val Gly Ser Asn Ala 670 675 680	2064
TAC GAA ATC CGC GAC CGC AAG ACC GGT GCC GTG AAG TGG ACC GCC TCG Tyr Glu Ile Arg Asp Arg Lys Thr Gly Ala Val Lys Trp Thr Ala Ser 685 690 700	2112
CGG GTG GAT CTG GTA TTT GGT TCC AAC TCG CTA CTG CGC TCT TAC GCA Arg Val Asp Leu Val Phe Gly Ser Asn Ser Leu Leu Arg Ser Tyr Ala 705 710 715 720	2160
GAA GTG TAC GCC CAG GAC GAT AAC GGC GAG AAG TTC GTC AGA GAC TTC Glu Val Tyr Ala Gln Asp Asp Asn Gly Glu Lys Phe Val Arg Asp Phe 725 730 735	2208
GTC GCC GCC TGG ACC AAA GTG ATG AAC GCC GAC CGT TTC GAC GTC GCG Val Ala Ala Trp Thr Lys Val Met Asn Ala Asp Arg Phe Asp Val Ala 740 745 750	2256
TCG TAA Ser	2262

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 753 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Polypeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Asn Ala Ser Ala Asp Asp Leu His Ser Ser Leu Gln Gln Arg 5 10 15
Cys Arg Ala Phe Val Pro Leu Val Ser Pro Arg His Arg Ala Ile Arg 20 25 30
Glu Arg Ala Met Ser Gly Lys Cys Pro Val Met His Gly Gly Asn Thr 35 40 45

Ser Thr Gly Thr Ser Asn Lys Asp Trp Trp Pro Glu Gly Leu Asn Leu
 50 55 60
 Asp Ile Leu His Gln Gln Asp Arg Lys Ser Asp Pro Met Asp Pro Asp
 65 70 75 80
 Phe Asn Tyr Arg Glu Glu Val Arg Lys Leu Asp Phe Asp Ala Leu Lys
 85 90 95
 Lys Asp Val His Ala Leu Met Thr Asp Ser Gln Glu Trp Trp Pro Ala
 100 105 110
 Asp Trp Gly His Tyr Gly Gly Leu Met Ile Arg Met Ala Trp His Ser
 115 120 125
 Ala Gly Thr Tyr Arg Ile Ala Asp Gly Arg Gly Gly Gly Thr Gly
 130 135 140
 Ser Gln Arg Phe Ala Pro Leu Asn Ser Trp Pro Asp Asn Val Ser Leu
 145 150 155 160
 Asp Lys Ala Arg Arg Leu Leu Trp Pro Ile Lys Lys Lys Tyr Gly Asn
 165 170 175
 Lys Ile Ser Trp Ala Asp Leu Met Ile Leu Ala Gly Thr Val Ala Tyr
 180 185 190
 Glu Ser Met Gly Leu Pro Ala Tyr Gly Phe Ser Phe Gly Arg Val Asp
 195 200 205
 Ile Trp Glu Pro Glu Lys Asp Ile Tyr Trp Gly Asp Glu Lys Glu Trp
 210 215 220
 Leu Ala Pro Ser Asp Glu Arg Tyr Gly Asp Val Asn Lys Pro Glu Thr
 225 230 235 240
 Met Glu Asn Pro Leu Ala Ala Val Gln Met Gly Leu Ile Tyr Val Asn
 245 250 255
 Pro Glu Gly Val Asn Gly His Pro Asp Pro Leu Arg Thr Ala Gln Gln
 260 265 270
 Val Leu Glu Thr Phe Ala Arg Met Ala Met Asn Asp Glu Lys Thr Ala
 275 280 285
 Ala Leu Thr Ala Gly Gly His Thr Val Gly Asn Cys His Gly Asn Gly
 290 295 300
 Asn Ala Ser Ala Leu Ala Pro Asp Pro Lys Ala Ser Asp Val Glu Asn
 305 310 315 320
 Gln Gly Leu Gly Trp Gly Asn Pro Asn Met Gln Gly Lys Ala Ser Asn
 325 330 335
 Ala Val Thr Ser Gly Ile Glu Gly Ala Trp Thr Thr Asn Pro Thr Lys
 340 345 350
 Phe Asp Met Gly Tyr Phe Asp Leu Leu Phe Gly Tyr Asn Trp Glu Leu
 355 360 365
 Lys Lys Ser Pro Ala Gly Ala His His Trp Glu Pro Ile Asp Ile Lys
 370 375 380
 Lys Glu Asn Lys Pro Val Asp Ala Ser Asp Pro Ser Ile Arg His Asn
 385 390 395 400

Pro Ile Met Thr Asp Ala Asp Met Ala Ile Lys Val Asn Pro Thr Tyr
 405 410 415
 Arg Ala Ile Cys Glu Lys Phe Met Ala Asp Pro Glu Tyr Phe Lys Lys
 420 425 430
 Thr Phe Ala Lys Ala Trp Phe Lys Leu Thr His Arg Asp Leu Gly Pro
 435 440 445
 Lys Ser Arg Tyr Ile Gly Pro Glu Val Pro Ala Glu Asp Leu Ile Trp
 450 455 460
 Gln Asp Pro Ile Pro Ala Gly Asn Thr Asp Tyr Cys Glu Glu Val Val
 465 470 475 480
 Lys Gln Lys Ile Ala Gln Ser Gly Leu Ser Ile Ser Glu Met Val Ser
 485 490 495
 Thr Ala Trp Asp Ser Ala Arg Thr Tyr Arg Gly Ser Asp Met Arg Gly
 500 505 510
 Gly Ala Asn Gly Ala Arg Ile Arg Leu Ala Pro Gln Asn Glu Trp Gln
 515 520 525
 Gly Asn Glu Pro Glu Arg Leu Ala Lys Val Leu Ser Val Tyr Glu Gln
 530 535 540
 Ile Ser Ala Asp Thr Gly Ala Ser Ile Ala Asp Val Ile Val Leu Ala
 545 550 555 560
 Gly Ser Val Gly Ile Glu Lys Ala Ala Lys Ala Ala Gly Tyr Asp Val
 565 570 575
 Arg Val Pro Phe Leu Lys Gly Arg Gly Asp Ala Thr Ala Glu Met Thr
 580 585 590
 Asp Ala Asp Ser Phe Ala Pro Leu Glu Pro Leu Ala Asp Gly Phe Arg
 595 600 605
 Asn Trp Gln Lys Lys Glu Tyr Val Val Lys Pro Glu Glu Met Leu Leu
 610 615 620
 Asp Arg Ala Gln Leu Met Gly Leu Thr Gly Pro Glu Met Thr Val Leu
 625 630 635 640
 Leu Gly Gly Met Arg Val Leu Gly Thr Asn Tyr Gly Gly Thr Lys His
 645 650 655
 Gly Val Phe Thr Asp Cys Glu Gly Gln Leu Thr Asn Asp Phe Phe Val
 660 665 670
 Asn Leu Thr Asp Met Gly Asn Ser Trp Lys Pro Val Gly Ser Asn Ala
 675 680 685
 Tyr Glu Ile Arg Asp Arg Lys Thr Gly Ala Val Lys Trp Thr Ala Ser
 690 695 700
 Arg Val Asp Leu Val Phe Gly Ser Asn Ser Leu Leu Arg Ser Tyr Ala
 705 710 715 720
 Glu Val Tyr Ala Gln Asp Asp Asn Gly Glu Lys Phe Val Arg Asp Phe
 725 730 735
 Val Ala Ala Trp Thr Lys Val Met Asn Ala Asp Arg Phe Asp Val Ala
 740 745 750
 Ser

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 2238 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATG GAA AAT CAC AAA CAC TCA GGA TCT TCT ACG TAT AAC ACA AAC ACT	48
Met Glu Asn His Lys His Ser Gly Ser Ser Thr Tyr Asn Thr Asn Thr	
5 10 15	
GGC GGA AAA TGC CCT TTT ACC GGA GGT TCG CTT AAG CAA AGT GCA GGT	96
Gly Gly Lys Cys Pro Phe Thr Gly Gly Ser Leu Lys Gln Ser Ala Gly	
20 25 30	
GGC GGC ACC AAA AAC AGG GAT TGG TGG CCC AAC ATG CTC AAC CTC GGC	144
Gly Gly Thr Lys Asn Arg Asp Trp Trp Pro Asn Met Leu Asn Leu Gly	
35 40 45	
ATC TTA CGC CAA CAT TCA TCG CTA TCG GAC CCA AAC GAC CCG GAT TTT	192
Ile Leu Arg Gln His Ser Ser Leu Ser Asp Pro Asn Asp Pro Asp Phe	
50 55 60	
GAC TAT GCC GAA GAG TTT AAG AAG CTA GAT CTG GCA GCG GTT AAA AAG	240
Asp Tyr Ala Glu Glu Phe Lys Lys Leu Asp Leu Ala Ala Val Lys Lys	
65 70 75 80	
GAC CTG GCA GCG CTA ATG ACA GAT TCA CAG GAC TGG TGG CCA GCA GAT	288
Asp Leu Ala Ala Leu Met Thr Asp Ser Gln Asp Trp Trp Pro Ala Asp	
85 90 95	
TAC GGT CAT TAT GGC CCC TTC TTT ATA CGC ATG GCG TGG CAC AGC GCC	336
Tyr Gly His Tyr Gly Pro Phe Phe Ile Arg Met Ala Trp His Ser Ala	
100 105 110	
GGC ACC TAC CGT ATC GGT GAT GGC CGT GGT GGC GGT GGC TCC GGC TCA	384
Gly Thr Tyr Arg Ile Gly Asp Gly Arg Gly Gly Gly Gly Ser Gly Ser	
115 120 125	
CAG CGC TTC GCG CCT CTC AAT AGC TGG CCA GAC AAT GCC AAT CTG GAT	432
Gln Arg Phe Ala Pro Leu Asn Ser Trp Pro Asp Asn Ala Asn Leu Asp	
130 135 140	
AAA GCA CGC TTG CTT CTT TGG CCC ATC AAA CAA AAA TAC GGT CGA AAA	480
Lys Ala Arg Leu Leu Leu Trp Pro Ile Lys Gln Lys Tyr Gly Arg Lys	
145 150 155 160	
ATC TCC TGG GCG GAT CTA ATG ATA CTC ACA GGA AAC GTA GCT CTG GAA	528
Ile Ser Trp Ala Asp Leu Met Ile Leu Thr Gly Asn Val Ala Leu Glu	
165 170 175	
ACT ATG GGC TTT AAA ACT TTT GGT TTT GCA GGT GGC AGA GCA GAT GTA	576
Thr Met Gly Phe Lys Thr Phe Gly Phe Ala Gly Gly Arg Ala Asp Val	
180 185 190	
TGG GAG CCT GAA GAA GAT GTA TAC TGG GGA GCA GAA ACC GAA TGG CTG	624
Trp Glu Pro Glu Glu Asp Val Tyr Trp Gly Ala Glu Thr Glu Trp Leu	
195 200 205	

GGA GAC AAG CGC TAT GAA GGT GAC CGA GAG CTC GAA AAT CCC CTG GGA Gly Asp Lys Arg Tyr Glu Gly Asp Arg Glu Leu Glu Asn Pro Leu Gly 210 215 220	672
GCC GTA CAA ATG GGA CTC ATC TAT GTA AAC CCC GAA GGA CCC AAC GGC Ala Val Gln Met Gly Leu Ile Tyr Val Asn Pro Glu Gly Pro Asn Gly 225 230 235 240	720
AAG CCA GAC CCT ATC GCT GCT GCG CGT GAT ATT CGT GAG ACT TTT GGC Lys Pro Asp Pro Ile Ala Ala Ala Arg Asp Ile Arg Glu Thr Phe Gly 245 250 255	768
CGA ATG GCA ATG AAT GAC GAA GAA ACC GTG GCT CTC ATA GCG GGT GGA Arg Met Ala Met Asn Asp Glu Glu Thr Val Ala Leu Ile Ala Gly Gly 260 265 270	816
CAC ACC TTC GGA AAA ACC CAT GGT GCT GCC GAT GCG GAG AAA TAT GTG His Thr Phe Gly Lys Thr His Gly Ala Ala Asp Ala Glu Lys Tyr Val 275 280 285	864
GGC CGA GAG CCT GCC GCC GCA GGT ATT GAA GAA ATG AGC CTG GGG TGG Gly Arg Glu Pro Ala Ala Ala Gly Ile Glu Glu Met Ser Leu Gly Trp 290 295 300	912
AAA AAC ACC TAC GGC ACC GGA CAC GGT GCG GAT ACC ATC ACC AGT GGA Lys Asn Thr Tyr Gly Thr Gly His Gly Ala Asp Thr Ile Thr Ser Gly 305 310 315 320	960
CTA GAA GGC GCC TGG ACC AAG ACC CCT ACT CAA TGG AGC AAT AAC TTT Leu Glu Gly Ala Trp Thr Lys Thr Pro Thr Gln Trp Ser Asn Asn Phe 325 330 335	1008
TTT GAA AAC CTC TTT GGT TAC GAG TGG GAG CTT ACC AAA AGT CCA GCT Phe Glu Asn Leu Phe Gly Tyr Glu Trp Glu Leu Thr Lys Ser Pro Ala 340 345 350	1056
GGA GCT TAT CAG TGG AAA CCA AAA GAC GGT GCC GGG GCT GGC ACC ATA Gly Ala Tyr Gln Trp Lys Pro Lys Asp Gly Ala Gly Ala Gly Thr Ile 355 360 365	1104
CCG GAT GCA CAT GAT CCC AGC AAG TCG CAC GCT CCA TTT ATG CTC ACT Pro Asp Ala His Asp Pro Ser Lys Ser His Ala Pro Phe Met Leu Thr 370 375 380	1152
ACG GAC CTG GCG CTG CGC ATG GAC CCT GAT TAC GAA AAA ATT TCT CGA Thr Asp Leu Ala Leu Arg Met Asp Pro Asp Tyr Glu Lys Ile Ser Arg 385 390 395 400	1200
CGG TAC TAT GAA AAC CCT GAT GAG TTT GCA GAT GCT TTC GCG AAA GCA Arg Tyr Tyr Glu Asn Pro Asp Glu Phe Ala Asp Ala Phe Ala Lys Ala 405 410 415	1248
TGG TAC AAA CTG ACA CAC AGA GAT ATG GGA CCA AAG GTG CGC TAC CTG Trp Tyr Lys Leu Thr His Arg Asp Met Gly Pro Lys Val Arg Tyr Leu 420 425 430	1296
GGA CCA GAA GTG CCT CAG GAA GAC CTC ATC TGG CAA GAC CCT ATA CCA Gly Pro Glu Val Pro Gln Glu Asp Leu Ile Trp Gln Asp Pro Ile Pro 435 440 445	1344
GAT GTA AGC CAT CCT CTT GTA GAC GAA AAC GAT ATT GAA GGC CTA AAA Asp Val Ser His Pro Leu Val Asp Glu Asn Asp Ile Glu Gly Leu Lys 450 455 460	1392
GCC AAA ATC CTG GAA TCG GGA CTG ACG GTA AGC GAG CTG GTA AGC ACG Ala Lys Ile Leu Glu Ser Gly Leu Thr Val Ser Glu Leu Val Ser Thr 465 470 475 480	1440

GCA TGG GCT TCT GCA TCT ACT TTT AGA AAC TCT GAC AAG CGC GGC GGT Ala Trp Ala Ser 485 Ala Ser Thr Phe Arg Asn Ser Asp Lys Arg Gly Gly 500	1488
GCC AAC GGT GCA CGT ATA CGA CTG GCC CCA CAA AAA GAC TGG GAA GTA Ala Asn Gly 505 Ala Arg Ile Arg Leu Ala Pro Gln Lys Asp Trp Glu Val 515	1536
AAC AAC CCT CAG CAA CTT GCC AGG GTA CTC AAA ACA CTA GAA GGT ATC Asn Asn Pro 520 Gln Gln Leu Ala Arg Val Leu Lys Thr 530 Leu Glu Gly Ile	1584
CAG GAG GAC TTT AAC CAG GCG CAA TCA GAT AAC AAA GCA GTA TCG TTG Gln Glu Asp 535 Phe Asn Gln Ala Gln Ser Asp Asn Lys 545 Ala Val Ser Leu	1632
GCC GAC CTG ATT GTG CTG GCC GGC TGT GCG GGT GTA GAA AAA GCT GCA Ala Asp Leu Ile Val Leu Ala Gly Cys Ala 550 Gly Val Glu Lys Ala 565	1680
AAA GAT GCT GGC CAT GAG GTG CAG GTG CCT TTC AAC CCG GGA CGA GCG Lys Asp Ala Gly 570 His Glu Val Gln Val Pro Phe Asn Pro Gly Arg Ala 580	1728
GAT GCC ACC GCT GAG CAA ACC GAT GTG GAA GCT TTC GAA GCA CTA GAG Asp Ala Thr 585 Ala Glu Gln Thr Asp Val Glu Ala Phe Glu Ala Leu Glu 595	1776
CCA GCG GCT GAC GGC TTT AGA AAC TAC ATT AAA CCG GAG CAT AAA GTA Pro Ala Ala Asp 600 Gly Phe Arg Asn Tyr Ile Lys Pro Glu His Lys Val 610	1824
TCC GCT GAG GAA ATG CTC GTA GAC CGG GCG CAG CTT CTG TCG CTT TCG Ser Ala Glu Glu Met Leu Val Asp Arg Ala Gln Leu Leu Ser Leu Ser 625	1872
GCA CCA GAA ATG ACT GCT TTG GTA GGC GGT ATG CGT GTA CTG GGC ACC Ala Pro Glu Met Thr Ala Leu Val Gly Gly Met Arg Val Leu Gly Thr 630 635 640 645	1920
AAC TAC GAC GGT TCG CAG CAT GGA GTG TTT ACA AAT AAG CCG GGT CAG Asn Tyr Asp 650 Gly Ser Gln His Gly Val Phe Thr Asn Lys Pro Gly Gln 660	1968
CTA TCC AAT GAC TTC TTT GTA AAC CTG CTA GAC CTC AAC ACT AAA TGG Leu Ser Asn Asp 665 Phe Phe Val Asn Leu Leu Asp Leu Asn Thr Lys Trp 670 675	2016
CGA GCC AGC GAT GAA TCA GAC AAA GTT TTT GAA GGC AGA GAC TTC AAA Arg Ala Ser Asp 680 Glu Ser Asp Lys Val Phe Glu Gly Arg Asp Phe Lys 685 690	2064
ACT GGC GAA GTA AAG TGG AGT GGC ACC CGG GTA GAC CTG ATC TTC GGA Thr Gly Glu Val Lys Trp 695 Ser Gly Thr Arg Val Asp Leu Ile Phe Gly 700 710	2112
TCC AAT TCC GAG CTA AGA GCC CTC GCA GAA GTG TAC GGC TGT GCA GAT Ser Asn Ser Glu Leu Arg Ala Leu Ala Glu Val Tyr Gly Cys Ala Asp 715 720 725 730	2160
TCT GAA GAA AAG TTT GTT AAA GAT TTT GTG AAG GCC TGG GCC AAA GTA Ser Glu Glu Lys 735 Phe Val Lys Asp Phe Val Lys Ala Trp Ala Lys Val 740 745	2208
ATG GAC CTG GAC CGG TTT GAT CTG AAA TAA Met Asp Leu Asp Arg Phe Asp Leu Lys 750 755	2238

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 745 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Asn His Lys His Ser Gly Ser Ser Thr Tyr Asn Thr Asn Thr
5 10 15
Gly Gly Lys Cys Pro Phe Thr Gly Gly Ser Leu Lys Gln Ser Ala Gly
20 25 30
Gly Gly Thr Lys Asn Arg Asp Trp Trp Pro Asn Met Leu Asn Leu Gly
35 40 45
Ile Leu Arg Gln His Ser Ser Leu Ser Asp Pro Asn Asp Pro Asp Phe
50 55 60
Asp Tyr Ala Glu Glu Phe Lys Lys Leu Asp Leu Ala Ala Val Lys Lys
65 70 75 80
Asp Leu Ala Ala Leu Met Thr Asp Ser Gln Asp Trp Trp Pro Ala Asp
85 90 95
Tyr Gly His Tyr Gly Pro Phe Phe Ile Arg Met Ala Trp His Ser Ala
100 105 110
Gly Thr Tyr Arg Ile Gly Asp Gly Arg Gly Gly Gly Gly Ser Gly Ser
115 120 125
Gln Arg Phe Ala Pro Leu Asn Ser Trp Pro Asp Asn Ala Asn Leu Asp
130 135 140
Lys Ala Arg Leu Leu Leu Trp Pro Ile Lys Gln Lys Tyr Gly Arg Lys
145 150 155 160
Ile Ser Trp Ala Asp Leu Met Ile Leu Thr Gly Asn Val Ala Leu Glu
165 170 175
Thr Met Gly Phe Lys Thr Phe Gly Phe Ala Gly Gly Arg Ala Asp Val
180 185 190
Trp Glu Pro Glu Glu Asp Val Tyr Trp Gly Ala Glu Thr Glu Trp Leu
195 200 205
Gly Asp Lys Arg Tyr Glu Gly Asp Arg Glu Leu Glu Asn Pro Leu Gly
210 215 220
Ala Val Gln Met Gly Leu Ile Tyr Val Asn Pro Glu Gly Pro Asn Gly
225 230 235 240
Lys Pro Asp Pro Ile Ala Ala Ala Arg Asp Ile Arg Glu Thr Phe Gly
245 250 255
Arg Met Ala Met Asn Asp Glu Glu Thr Val Ala Leu Ile Ala Gly Gly
260 265 270
His Thr Phe Gly Lys Thr His Gly Ala Ala Asp Ala Glu Lys Tyr Val
275 280 285

Gly Arg Glu Pro Ala Ala Ala Gly Ile Glu Glu Met Ser Leu Gly Trp
 290 295 300
 Lys Asn Thr Tyr Gly Thr Gly His Gly Ala Asp Thr Ile Thr Ser Gly
 305 310 315 320
 Leu Glu Gly Ala Trp Thr Lys Thr Pro Thr Gln Trp Ser Asn Asn Phe
 325 330 335
 Phe Glu Asn Leu Phe Gly Tyr Glu Trp Glu Leu Thr Lys Ser Pro Ala
 340 345 350
 Gly Ala Tyr Gln Trp Lys Pro Lys Asp Gly Ala Gly Ala Gly Thr Ile
 355 360 365
 Pro Asp Ala His Asp Pro Ser Lys Ser His Ala Pro Phe Met Leu Thr
 370 375 380
 Thr Asp Leu Ala Leu Arg Met Asp Pro Asp Tyr Glu Lys Ile Ser Arg
 385 390 395 400
 Arg Tyr Tyr Glu Asn Pro Asp Glu Phe Ala Asp Ala Phe Ala Lys Ala
 405 410 415
 Trp Tyr Lys Leu Thr His Arg Asp Met Gly Pro Lys Val Arg Tyr Leu
 420 425 430
 Gly Pro Glu Val Pro Gln Glu Asp Leu Ile Trp Gln Asp Pro Ile Pro
 435 440 445
 Asp Val Ser His Pro Leu Val Asp Glu Asn Asp Ile Glu Gly Leu Lys
 450 455 460
 Ala Lys Ile Leu Glu Ser Gly Leu Thr Val Ser Glu Leu Val Ser Thr
 465 470 475 480
 Ala Trp Ala Ser Ala Ser Thr Phe Arg Asn Ser Asp Lys Arg Gly Gly
 485 490 495
 Ala Asn Gly Ala Arg Ile Arg Leu Ala Pro Gln Lys Asp Trp Glu Val
 500 505 510
 Asn Asn Pro Gln Gln Leu Ala Arg Val Leu Lys Thr Leu Glu Gly Ile
 515 520 525
 Gln Glu Asp Phe Asn Gln Ala Gln Ser Asp Asn Lys Ala Val Ser Leu
 530 535 540
 Ala Asp Leu Ile Val Leu Ala Gly Cys Ala Gly Val Glu Lys Ala Ala
 545 550 555 560
 Lys Asp Ala Gly His Glu Val Gln Val Pro Phe Asn Pro Gly Arg Ala
 565 570 575
 Asp Ala Thr Ala Glu Gln Thr Asp Val Glu Ala Phe Glu Ala Leu Glu
 580 585 590
 Pro Ala Ala Asp Gly Phe Arg Asn Tyr Ile Lys Pro Glu His Lys Val
 595 600 605
 Ser Ala Glu Glu Met Leu Val Asp Arg Ala Gln Leu Leu Ser Leu Ser
 610 615 620
 Ala Pro Glu Met Thr Ala Leu Val Gly Gly Met Arg Val Leu Gly Thr
 625 630 635 640

Asn	Tyr	Asp	Gly	Ser	Gln	His	Gly	Val	Phe	Thr	Asn	Lys	Pro	Gly	Gln
				645					650					655	
Leu	Ser	Asn	Asp	Phe	Phe	Val	Asn	Leu	Leu	Asp	Leu	Asn	Thr	Lys	Trp
			660					665					670		
Arg	Ala	Ser	Asp	Glu	Ser	Asp	Lys	Val	Phe	Glu	Gly	Arg	Asp	Phe	Lys
		675					680					685			
Thr	Gly	Glu	Val	Lys	Trp	Ser	Gly	Thr	Arg	Val	Asp	Leu	Ile	Phe	Gly
	690					695					700				
Ser	Asn	Ser	Glu	Leu	Arg	Ala	Leu	Ala	Glu	Val	Tyr	Gly	Cys	Ala	Asp
705					710					715					720
Ser	Glu	Glu	Lys	Phe	Val	Lys	Asp	Phe	Val	Lys	Ala	Trp	Ala	Lys	Val
				725					730					735	
Met	Asp	Leu	Asp	Arg	Phe	Asp	Leu	Lys							
			740					745							

N:\R\251010.SEQ

09922185-080201